



OIFE

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/981,430

TIME: 15:16:57

Input Set : N:\Crf3\RULE60\09981430.txt

Output Set: N:\CRF3\01152002\I981430.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Genencor International, Inc.

7 (ii) TITLE OF INVENTION: ESTERASE ENZYMES, DNA ENCODING

8 ESTERASE ENZYMES AND VECTORS AND HOST CELLS INCORPORATING

SAME

10 (iii) NUMBER OF SEQUENCES: 35

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Genencor International, Inc.

14 (B) STREET: 925 Page Mill Road

15 (C) CITY: Palo Alto

16 (D) STATE: CA

17 (E) COUNTRY: USA

18 (F) ZIP: 94304-1013

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Diskette

22 (B) COMPUTER: IBM Compatible

23 (C) OPERATING SYSTEM: DOS

24 (D) SOFTWARE: FastSEQ for Windows Version 2.0

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/09/981,430

C--> 28 (B) FILING DATE: 15-Oct-2001

30 (vii) PRIOR APPLICATION DATA:

31 (A) APPLICATION NUMBER: 08/952,445

32 (B) FILING DATE:

34 (viii) ATTORNEY/AGENT INFORMATION:

35 (A) NAME: Stone, Christopher L.

36 (B) REGISTRATION NUMBER: 35,696

37 (C) REFERENCE/DOCKET NUMBER: GC362-2-US

39 (ix) TELECOMMUNICATION INFORMATION:

40 (A) TELEPHONE: 650-846-7555

41 (B) TELEFAX: 650-845-6504

44 (2) INFORMATION FOR SEQ ID NO: 1:

46 (i) SEQUENCE CHARACTERISTICS:

47 (A) LENGTH: 32 amino acids

48 (B) TYPE: amino acid

49 (C) STRANDEDNESS: single

50 (D) TOPOLOGY: linear

52 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

54 Ala Ser Thr Gln Gly Ile Ser Glu Asp Leu Tyr Ser Arg Leu Val Glu

55 1 5 10 15

W--> 56 Met Ala Thr Ile Ser Gln Ala Ala Tyr Xaa Asp Leu Leu Asn Ile Pro

57 20 25 30

59 (2) INFORMATION FOR SEQ ID NO: 2:

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/981,430

DATE: 01/15/2002

TIME: 15:16:57

Input Set : N:\Crf3\RULE60\09981430.txt

Output Set: N:\CRF3\01152002\I981430.raw

```

61      (i) SEQUENCE CHARACTERISTICS:
62          (A) LENGTH: 8 amino acids
63          (B) TYPE: amino acid
64          (C) STRANDEDNESS: single
65          (D) TOPOLOGY: linear
67      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
W--> 69  Xaa Thr Val Gly Phe Gly Pro Tyr
70      1          5
73  (2) INFORMATION FOR SEQ ID NO: 3:
75      (i) SEQUENCE CHARACTERISTICS:
76          (A) LENGTH: 9 amino acids
77          (B) TYPE: amino acid
78          (C) STRANDEDNESS: single
79          (D) TOPOLOGY: linear
81      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
W--> 83  Phe Gly Leu His Leu Xaa Gln Xaa Met
84      1          5
87  (2) INFORMATION FOR SEQ ID NO: 4:
89      (i) SEQUENCE CHARACTERISTICS:
90          (A) LENGTH: 8 amino acids
91          (B) TYPE: amino acid
92          (C) STRANDEDNESS: single
93          (D) TOPOLOGY: linear
95      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
W--> 97  Xaa Ile Ser Glu Asp Leu Tyr Ser
98      1          5
101  (2) INFORMATION FOR SEQ ID NO: 5:
103      (i) SEQUENCE CHARACTERISTICS:
104          (A) LENGTH: 9 amino acids
105          (B) TYPE: amino acid
106          (C) STRANDEDNESS: single
107          (D) TOPOLOGY: linear
109      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
111  Tyr Ile Gly Trp Ser Phe Tyr Asn Ala
112      1          5
115  (2) INFORMATION FOR SEQ ID NO: 6:
117      (i) SEQUENCE CHARACTERISTICS:
118          (A) LENGTH: 10 amino acids
119          (B) TYPE: amino acid
120          (C) STRANDEDNESS: single
121          (D) TOPOLOGY: linear
122      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
W--> 124  Gly Ile Ser Glu Asp Leu Tyr Xaa Xaa Gln
125      1          5          10
128  (2) INFORMATION FOR SEQ ID NO: 7:
130      (i) SEQUENCE CHARACTERISTICS:
131          (A) LENGTH: 10 amino acids
132          (B) TYPE: amino acid
133          (C) STRANDEDNESS: single

```

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/981,430

TIME: 15:16:57

Input Set : N:\Crf3\RULE60\09981430.txt

Output Set: N:\CRF3\01152002\I981430.raw

```

134          (D) TOPOLOGY: linear
136      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
W--> 138  Xaa Ile Ser Glu Ser Leu Tyr Xaa Xaa Arg
139      1          5          10
142  (2) INFORMATION FOR SEQ ID NO: 8:
144      (i) SEQUENCE CHARACTERISTICS:
145          (A) LENGTH: 7 amino acids
146          (B) TYPE: amino acid
147          (C) STRANDEDNESS: single
148          (D) TOPOLOGY: linear
150      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
152  Gly Ile Ser Glu Asp Leu Tyr
153      1          5
156  (2) INFORMATION FOR SEQ ID NO: 9:
158      (i) SEQUENCE CHARACTERISTICS:
159          (A) LENGTH: 7 amino acids
160          (B) TYPE: amino acid
161          (C) STRANDEDNESS: single
162          (D) TOPOLOGY: linear
164      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
166  Leu Glu Pro Pro Tyr Thr Gly
167      1          5
170  (2) INFORMATION FOR SEQ ID NO: 10:
172      (i) SEQUENCE CHARACTERISTICS:
173          (A) LENGTH: 14 amino acids
174          (B) TYPE: amino acid
175          (C) STRANDEDNESS: single
176          (D) TOPOLOGY: linear
178      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
W--> 180  Xaa Ala Asn Asp Gly Ile Pro Asn Leu Pro Pro Val Glu Gln
181      1          5          10
183  (2) INFORMATION FOR SEQ ID NO: 11:
185      (i) SEQUENCE CHARACTERISTICS:
186          (A) LENGTH: 8 amino acids
187          (B) TYPE: amino acid
188          (C) STRANDEDNESS: single
189          (D) TOPOLOGY: linear
191      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
193  Tyr Pro Asp Tyr Ala Leu Tyr Lys
194      1          5
197  (2) INFORMATION FOR SEQ ID NO: 12:
199      (i) SEQUENCE CHARACTERISTICS:
200          (A) LENGTH: 22 base pairs
201          (B) TYPE: nucleic acid
202          (C) STRANDEDNESS: single
203          (D) TOPOLOGY: linear
205      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
207  CGGGAATTCG CWSACCARGG AT
210  (2) INFORMATION FOR SEQ ID NO: 13:

```

22

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/981,430

TIME: 15:16:57

Input Set : N:\Crf3\RULE60\09981430.txt

Output Set: N:\CRF3\01152002\I981430.raw

```

212      (i) SEQUENCE CHARACTERISTICS:
213          (A) LENGTH: 32 amino acids
214          (B) TYPE: amino acid
215          (C) STRANDEDNESS: single
216          (D) TOPOLOGY: linear
219      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
221  Ala Ser Thr Gln Gly Ile Ser Glu Asp Leu Tyr Ser Arg Leu Val Glu
222   1          5          10          15
223  Met Ala Thr Ile Ser Gln Ala Ala Tyr Ala Asp Leu Leu Asn Ile Pro
224   20          25          30
227 (2) INFORMATION FOR SEQ ID NO: 14:
229      (i) SEQUENCE CHARACTERISTICS:
230          (A) LENGTH: 25 base pairs
231          (B) TYPE: nucleic acid
232          (C) STRANDEDNESS: single
233          (D) TOPOLOGY: linear
235      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
237  CGGGAATTCT AYTAYATHGG TGGGT
239 (2) INFORMATION FOR SEQ ID NO: 15:
241      (i) SEQUENCE CHARACTERISTICS:
242          (A) LENGTH: 16 amino acids
243          (B) TYPE: amino acid
244          (C) STRANDEDNESS: single
245          (D) TOPOLOGY: linear
247      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
249  Val His Gly Gly Tyr Tyr Ile Gly Trp Val Ser Val Gln Asp Gln Val
250   1          5          10          15
253 (2) INFORMATION FOR SEQ ID NO: 16:
255      (i) SEQUENCE CHARACTERISTICS:
256          (A) LENGTH: 25 base pairs
257          (B) TYPE: nucleic acid
258          (C) STRANDEDNESS: single
259          (D) TOPOLOGY: linear
261      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:
263  CGGGAATTCA CCCACCDATR TARTA
265 (2) INFORMATION FOR SEQ ID NO: 17:
267      (i) SEQUENCE CHARACTERISTICS:
268          (A) LENGTH: 16 amino acids
269          (B) TYPE: amino acid
270          (C) STRANDEDNESS: single
271          (D) TOPOLOGY: linear
273      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:
275  Val His Gly Gly Tyr Tyr Ile Gly Trp Val Ser Val Gln Asp Gln Val
276   1          5          10          15
279 (2) INFORMATION FOR SEQ ID NO: 18:
281      (i) SEQUENCE CHARACTERISTICS:
282          (A) LENGTH: 23 base pairs
283          (B) TYPE: nucleic acid
284          (C) STRANDEDNESS: single

```

25

25

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/981,430

TIME: 15:16:57

Input Set : N:\Crf3\RULE60\09981430.txt

Output Set: N:\CRF3\01152002\I981430.raw

```

285         (D) TOPOLOGY: linear
287     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:
289     CGGGAATTCT TGGATCCRTC RTT                                     23
292 (2) INFORMATION FOR SEQ ID NO: 19:
294     (i) SEQUENCE CHARACTERISTICS:
295         (A) LENGTH: 27 amino acids
296         (B) TYPE: amino acid
297         (C) STRANDEDNESS: single
298         (D) TOPOLOGY: linear
300     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:
302     Thr Asp Ala Phe Gln Ala Ser Ser Pro Asp Thr Thr Gln Tyr Phe Arg
303     1             5             10             15
304     Val Thr His Ala Asn Asp Gly Ile Pro Asn Leu
305             20             25
307 (2) INFORMATION FOR SEQ ID NO: 20:
309     (i) SEQUENCE CHARACTERISTICS:
310         (A) LENGTH: 24 base pairs
311         (B) TYPE: nucleic acid
312         (C) STRANDEDNESS: single
313         (D) TOPOLOGY: linear
315     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:
317     CGGGAATTCA TCCRTCRTTG CRTG                                     24
320 (2) INFORMATION FOR SEQ ID NO: 21:
322     (i) SEQUENCE CHARACTERISTICS:
323         (A) LENGTH: 27 amino acids
324         (B) TYPE: amino acid
325         (C) STRANDEDNESS: single
326         (D) TOPOLOGY: linear
328     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:
330     Thr Asp Ala Phe Gln Ala Ser Ser Pro Asp Thr Thr Gln Tyr Phe Arg
331     1             5             10             15
332     Val Thr His Ala Asn Asp Gly Ile Pro Asn Leu
333             20             25
336 (2) INFORMATION FOR SEQ ID NO: 22:
338     (i) SEQUENCE CHARACTERISTICS:
339         (A) LENGTH: 27 base pairs
340         (B) TYPE: nucleic acid
341         (C) STRANDEDNESS: single
342         (D) TOPOLOGY: linear
344     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:
346     CGGGAATTTCG CYTGRAAGCR TCGTCAT                               27
349 (2) INFORMATION FOR SEQ ID NO: 23:
351     (i) SEQUENCE CHARACTERISTICS:
352         (A) LENGTH: 28 amino acids
353         (B) TYPE: amino acid
354         (C) STRANDEDNESS: single
355         (D) TOPOLOGY: linear
357     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:
359     Met Thr Asp Ala Phe Gln Ala Ser Ser Pro Asp Thr Thr Gln Tyr Phe

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/981,430

DATE: 01/15/2002

TIME: 15:16:58

Input Set : N:\Crf3\RULE60\09981430.txt

Output Set: N:\CRF3\01152002\I981430.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:56 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:69 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:83 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:97 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:124 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:138 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:180 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10